

AD_____

Award Number: DAMD17-97-1-7210

TITLE: Transcriptional Regulation of BRCA1

PRINCIPAL INVESTIGATOR: Shannon R. Payne, Ph.D.
Mary Claire King, Ph.D.

CONTRACTING ORGANIZATION: University of Washington
Seattle, Washington 98105-6613

REPORT DATE: October 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503</small>			<small>OMB NO. 074-0188</small>
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE October 1999	3. REPORT TYPE AND DATES COVERED Annual (15 Sep 98 - 14 Sep 99)	
4. TITLE AND SUBTITLE Transcriptional Regulation of BRCA1			5. FUNDING NUMBERS DAMD17-97-1-7210
6. AUTHOR(S) Shannon R. Payne, Ph.D. Mary Claire King, Ph.D.			8. PERFORMING ORGANIZATION REPORT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Washington Seattle, Washington 98105-6613 E-MAIL: spayne@u.washington.edu			
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			10. SPONSORING / MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES <p style="text-align: center;">This report contains colored photos</p>			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 Words) Germline mutations in <i>BRCA1</i> lead to an increased risk of breast and ovarian cancer, with loss of the second, normal allele critical to tumorigenesis. It was widely presumed that the cloning and characterization of genes involved in hereditary breast cancer would lead to a better understanding of the genesis of the more common non-inherited forms of breast cancer. The relative lack of somatic mutations found in <i>BRCA1</i> , however, has argued against its involvement in non-inherited breast cancer. Our research specifically addresses whether large genomic rearrangements are responsible for somatic inactivation of <i>BRCA1</i> . We characterized the types of large germline rearrangements that occur within the <i>BRCA1</i> region and investigated the contribution of two large germline rearrangements to breast cancer in a population-based series of breast cancer patients. In order to determine whether <i>BRCA1</i> is inactivated somatically by large rearrangement of <i>BRCA1</i> , we analyzed 92 breast carcinomas using loss of heterozygosity analysis, Long PCR, Southern analysis, and immunohistochemistry. Although two large germline rearrangements were detected in our series, no large somatic rearrangements were identified. As previously reported <i>BRCA1</i> protein was reduced in the majority of breast tumors of high histologic grade. Interestingly, reduced <i>BRCA1</i> protein in sporadic breast carcinomas was significantly associated with loss of the most 5' <i>BRCA1</i> intragenic marker.			
14. SUBJECT TERMS Breast Cancer			15. NUMBER OF PAGES 45
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

SRP Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

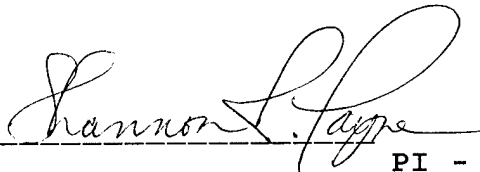
X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

X In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

November 1, 1999



PI - Signature

Date

Table of Contents
Predoctoral Fellowship Annual Report 1999
Shannon R. Payne
Grant DAMD17-97-1-7210

Front Cover	page 1
Standard Form 298	page 2
Foreword	page 3
Table of Contents	page 4
Introduction	pages 5-6
Body	pages 7-12
Tables 1-2	pages 13-14
Figures 1-6	pages 15-20
References	page 21-22
Key Accomplishments	page 23-24
Reportable Outcomes	page 25

Introduction

Germline mutations in *BRCA1* lead to an increased risk of breast and ovarian cancer, with loss of the second, wild type allele critical to tumorigenesis (1). More than 400 distinct *BRCA1* germline mutations have been identified to date (2). Mutation detection in *BRCA1* is complicated by the ubiquitous distribution of these mutations throughout the *BRCA1* coding region (2). As a result, most investigators have chosen to use a combination of techniques that rely on amplification of small, ~150-700 base pairs (bp) fragments from the *BRCA1* genomic region. This approach has worked relatively well for germline mutation screening of *BRCA1*, as most mutations are small insertions, deletions, or single bp substitutions that lead to premature protein truncation (2). However, mutation detection based on amplification of small fragments will not identify large-scale alterations of the *BRCA1* genomic region such as deletions, duplications, or inversions of hundreds of bp.

Of the breast and ovarian tumors screened by several independent groups only a few tumors with somatic mutation in *BRCA1* have been reported (1,3-5). All fourteen of the tumors harboring these somatic mutations exhibited LOH for the *BRCA1* region of chromosome 17.

Thus, *BRCA1* somatic mutations detectable by conventional PCR-based techniques are extremely rare in breast and ovarian tumors. These conventional screening procedures, however, neglect a class of mutations: large genomic rearrangements. Several large germline deletions and one large germline duplication in *BRCA1* have been reported (6-11). All of the large germline rearrangements were undetectable by conventional genomic DNA (gDNA) screening procedures. Large germline deletions appear to be a relatively common feature of inherited breast cancer in the Dutch population, due in large part to founder mutations (8).

Several features of the *BRCA1* locus provide clues as to the nature of large genomic rearrangements that might occur. *BRCA1* has one of the highest densities of Alu elements of genes deposited in GenBank (41.5% of 81 kb) (12). Large genomic rearrangements in human genes frequently are associated with Alu repetitive elements (13,14).

Another predisposing feature of the *BRCA1* locus is a tandem duplication involving the 5' region of *BRCA1* (15,16). The duplication includes the *BRCA1* promoter region and a neighboring gene, *NBR1*. The result of the duplication was the creation of a new gene, *NBR2*, with which human *BRCA1* shares a bi-directional promoter and the creation of a pseudo-copy of *BRCA1* situated in a similar head to head orientation with the *NBR1* gene approximately 30 kb upstream of the functional *BRCA1* locus (17). The presence of a pseudo-copy of *BRCA1* with >90% identity located so near on chromosome 17 further predisposes the functional *BRCA1* locus to inactivation by misaligned homologous recombination or gene conversion.

Project goals

Several possible interpretations could explain the paradox presented by the lack of somatic mutations in *BRCA1*. One formal explanation is that *BRCA1* is not involved in sporadic breast and ovarian cancer. It is possible that the high rate of LOH for the *BRCA1* region in sporadic tumors actually targets a nearby gene, and not *BRCA1*. A second explanation is an alternative mechanism for loss of the second *BRCA1* allele in sporadic tumors. A third possibility is that the lack of *BRCA1* mutations in sporadic tumors actually reflects limitations in current mutation screening procedures. Given the genomic organization of the *BRCA1* locus and the methods commonly used to screen for *BRCA1* mutations, it is conceivable that *BRCA1* somatic mutations in breast and ovarian cancer exist but have remained undetected by conventional mutation detection methods.

The research detailed in the following sections further elucidates the role of *BRCA1* in breast cancer by analyzing the genetic alterations that lead to inactivation of *BRCA1*. First, the characterization of large germline rearrangements identified in *BRCA1* and their contribution to breast cancer in a population-based series of breast cancer patients will be described. Second, the nature of somatic alterations identified from sporadic breast tumors and their relationship to *BRCA1* protein expression will be detailed. Finally, a rare variant allele of *BRCA1* and the potential significance of noncoding variation in the *BRCA1* genomic region will be described.

Body

Characterization of a large germline deletion in the BRCA1 region leading to breast and ovarian cancer predisposition

One large germline rearrangement was identified from Family 58 (also known as Family 5). The rearrangement consisted of a 1039 bp deletion with a 10 bp inverted duplication inserted at the breakpoint junction (g.12977 ins10 del1039). The research characterizing this deletion was described in my 1998 Annual Report and has been accepted for publication in *Genes, Chromosomes, & Cancer*.

No other large germline rearrangements were detected in the families or cell lines using either Long PCR or Southern analysis. Several restriction fragment length polymorphisms (RFLP's) were observed. The RFLP's identified and their observed frequencies are listed in Table 1.

Population-based investigation of two large germline rearrangements predisposing to breast and ovarian cancer

Next we investigated the contribution of two different large germline rearrangements to breast cancer in the American population. We analyzed one large deletion (the g.12977 ins10 del1039 mutation) and one large duplication (the g.43368ins6081 mutation) in a population-based series of American breast cancer patients unselected for family history (10,18,19). The g.12977 ins10 del1039 mutation was investigated in 242 white and 164 African-American breast cancer patients. No mutation carriers were detected in 406 cases. The g.43368ins6081 mutation was investigated in 220 white and 153 African-American breast cancer patients. One mutation carrier (CBCS 950749) was detected in 373 cases.

The g.43368ins6081 mutation was previously identified in four American families of mixed European descent with multiple cases of breast and/or ovarian cancer and one Portuguese family with three cases of breast cancer (10, Heather Mefford, pers. comm.). A founder effect for the g.43368ins6081 mutation was postulated based on the observation that four of the families shared the same haplotype at nine polymorphic markers within or flanking the *BRCA1* locus (10).

LOH of the BRCA1 region in sporadic breast carcinoma

The *BRCA1* genomic region investigated is shown in Figure 1. Of the paired breast normal and tumor samples informative for LOH, 50% (39/78) of invasive ductal tumors lost an allele at one or more markers within the *BRCA1* region. None of the 12 lobular carcinomas showed any allelic loss. Of the 56 tumors informative for the three *BRCA1* intragenic markers (D17S1323, D17S1322, D17S855), 36 tumors (64.3%) had lost at least one intragenic marker. Allelic loss at the three *BRCA1* intragenic markers as well as at the three flanking markers (D17S1320, D17S1326, D17S1325) was associated significantly with tumor grade. LOH was detected in 44% (4/9) of grade I tumors, 28% (11/40) of grade II tumors, and 63% (26/41) grade III tumors ($p < 0.005$). This association is consistent with previous reports of association between somatic LOH at the *BRCA1* locus and tumor grade (20-23). Estrogen receptor status was independently associated with LOH ($p = 0.046$) as reported previously (22,23). There was no independent association between LOH and tumor stage or age at diagnosis.

Microsatellite instability in an infiltrating ductal carcinoma

Case 12 was unique in that additional alleles appeared at four markers coincident with LOH at the remaining two markers. In order to test whether the additional alleles were due to microsatellite instability (MSI), we analyzed markers indicative of MSI for Case 12. These markers were shown previously to be diagnostic of a mismatch repair defect in tumor cells (24). Five of ten diagnostic markers were analyzed for Case 12: BAT26, BAT40, D2S123, D5S346, and Mfd15. Case 12 showed instability at four of the five markers tested. Only the BAT26 marker was stable in tumor gDNA. Thus, Case 12 showed 80% unstable loci among the first five markers.

Long PCR and Southern analysis of sporadic breast carcinomas

Of the 92 samples, 84 were amplified successfully for all 15 Long PCR primer pairs. However, no large somatic rearrangements were detected. Several RFLP's were observed in both normal and tumor gDNA from the cases. These are listed in Table 2. One large germline deletion in intron 15 of *BRCA1* was detected in Case 72.

In order to detect very large deletions, duplications, and rearrangements that would be missed by PCR, we also performed Southern analysis. As tumor DNA was limiting and because no large rearrangements were identified using Long PCR, Southern analysis was performed with only the 40 tumors that demonstrated LOH within the *BRCA1* region. The eight samples that were not screened fully by Long PCR were analyzed by Southern as well. One ~250 bp germline insertion in intron 7 of *BRCA1* was observed in Case 72. No large somatic rearrangements were detected.

BRCA1 protein expression and loss of heterozygosity

In collaboration with Rachel Gonzalez-Hernandez (Pathology Department, University of Washington), *BRCA1* protein expression in breast normal and tumor tissue sections was analyzed by immunohistochemistry (IHC) using monoclonal antibody MS110 as previously described (25). The normal and tumor breast tissues were scored on a scale of 0 to 4. Normal breast epithelium exhibited staining consistent with a score of 3. Reduced *BRCA1* protein (scores of 0 to 2) was observed in 72% (49/68) of the breast tumors and in the majority of tumors (82%, 28/34) with high histologic grade as compared to normal breast epithelial tissue ($p < 0.001$).

Reduced *BRCA1* protein expression was associated with genomic loss at D17S1323, the intragenic marker closest to the *BRCA1* transcription start site. Reduced *BRCA1* protein staining (scores of 0 to 2 on a scale of 0 to 4) occurred in 92% (11/12) of tumors with genomic loss at D17S1323 but in 55% (12/22) without genomic loss at this marker ($p < 0.02$). Genomic loss at marker D17S1325 was associated marginally with reduced *BRCA1* protein expression ($p = 0.05$), although the significance of this association is unclear. Marker D17S1325 is located ~300 kb upstream of *BRCA1*. Although marker D17S1326 lies between D17S1323 and D17S1325, LOH at D17S1326 is not associated with reduced *BRCA1* protein staining.

55% (12/22) of tumors in the series had reduced *BRCA1* protein staining (scores of 0 to 2) despite no detectable genomic loss of the *BRCA1* transcription start as indicated by marker D17S1323. Reduced *BRCA1* protein was observed in a high proportion (8/14, 57%) of the tumors that were informative at all three intragenic markers and fully

retained both alleles of *BRCA1*. In this subset of tumors, other mechanisms are likely responsible for *BRCA1* inactivation.

Characterization of a large germline deletion of unknown significance from Case 72

In most of the 217 alleles investigated, the general organization of the *BRCA1* region was consistent with the published sequence. More than 48 restriction sites over a region of more than 85 kb were analyzed in high-risk breast and/or ovarian cancer family members and sporadic breast cancer patients unselected for family history. In the course of analyzing the family and tumor DNA for large rearrangement of *BRCA1*, however, two large germline rearrangements of unknown significance were identified from a single allele in normal and tumor gDNA from Case 72. In addition, several small alterations were identified in the variant allele from Case 72 containing the two large germline rearrangements.

A large germline deletion of *BRCA1* was identified from Case 72. Using primer pair 10, designed to amplify a 10.6 kb product encompassing exons 15-18, variant restriction fragments were detected using four different restriction enzymes. The variant restriction fragments were present in equal intensity in gDNA from normal breast epithelial tissue of Case 72, but the variant bands were preferentially retained in gDNA prepared from an infiltrating ductal carcinoma of Case 72.

Based on the restriction digests, a minimal region of ~3800 bp (GenBank L78833 bp 54885-58706) containing a deletion of ~1000 bp was identified near *BRCA1* exon 16. Wild type and variant PCR products encompassing the deletion were cloned and sequenced. All clones of the variant PCR products contained the same SNP's that were not observed in clones from the wild type PCR products. However, all clones from the variant PCR products contained an intact exon 16, indicating that the breakpoint occurred in intron 15. Sequence of all variant clones revealed the simple deletion shown in Figure 2. Based on *BRCA1* genomic sequence L78833, the deletion is a 977 bp deletion of bp 56025-57001 from *BRCA1* intron 15. The notation for the deletion (26) is g.56024 del977 based on GenBank L78833. The 5' breakpoint for g.56024 del977 occurs in the 3' tail of an Alu-Sp element and the 3' breakpoint occurs 9 bp into the 5' head of an oppositely oriented Alu-Y element. Southern analysis using a probe to *BRCA1* exons 11-

24 confirmed the deletion and indicated that the deleted allele was retained in Case 72 tumor DNA.

Characterization of a large expansion of a compound microsatellite sequence in BRCA1 intron 7

Southern analysis of Hind III-digested Case 72 tumor gDNA revealed an extra band of ~4.3 kb. Further Southern analysis of Case 72 normal and tumor gDNA using multiple restriction enzymes indicated that a minimal region of ~3500 bp (GenBank L78833 bp 26,911-30394) including *BRCA1* exon 8 contained an insertion of ~250 bp (Figure 3). Overlapping PCR primer pairs were designed to span the region and tumor DNA from Case 72 was amplified. The exon 8 containing PCR product was of expected size, but a primer pair (ivs7b) amplifying a region just upstream of exon 8 (GenBank L78833 bp 27,656-28,733) yielded a product ~250 bp larger than expected from Case 72 tumor gDNA.

The variant ivs 7b product for Case 72 was gel-extracted and sequenced using internal primers. Multiple independent sequencing reactions revealed an expansion of a compound microsatellite located at GenBank L78833 bp ~28,087 to 28,174 (Figure 4). In the published genomic sequence the imperfect compound microsatellite GTN (AT)₄₋₈ is 87 bp in size. In the variant allele from Case 72, the compound microsatellite is 333 bp in size.

Other sequence alterations observed in the variant allele from Case 72

Several other alterations were observed in the variant allele retained in the tumor of Case 72 (Figure 5). Two single bp substitutions were observed in exon 16. One is a C to T transition at nucleotide (nt) 4801 of HSU14680. The substitution does not affect translation of the *BRCA1* message. This SNP has been observed before, although at a very low frequency. It was identified in one African-American patient in the Carolina Breast Cancer Study and is present in about 1% of the population (2). The second substitution is a A to G transition at nt 4956. The substitution results in a missense polymorphism at codon1613 (S1613G). This missense polymorphism has been observed before as well and is present in 32% of the population (27). Finally, a 24 bp insertion at

GenBank L78833 bp 57,211 was identified in the variant allele (Figure 6). The origin of the 24 bp insertion is unknown, but it contains a region of 13-20 bp identity with sequence from human chromosome 17 clone hRPK.268_F_2, human chromosome 6 PAC 271G9, and an unassigned human BAC RG191D16.

In summary, one large rearrangement identified from Case 72 involves germline deletion of 977 bp from *BRCA1* intron 15. The second large rearrangement involves germline expansion of a compound microsatellite sequence in *BRCA1* intron 7. While neither of the rearrangements alters coding regions of *BRCA1*, they are of interest because the allele that harbors both of these rearrangements is retained in an infiltrating ductal carcinoma from Case 72 while the wild-type allele of *BRCA1* is lost.

There is a 50% probability that the Case 72 variant allele of *BRCA1* is retained in the tumor randomly. In order to test whether the large rearrangements might affect processing of the *BRCA1* transcript from the rare allele, we analyzed splicing of exons flanking the intronic rearrangements. The large germline rearrangements do not affect splicing of *BRCA1* as determined by RT-PCR using a variety of primer pairs located in exons flanking the rearrangements. Additionally, we tested the variant allele for loss of transcript using single nucleotide polymorphisms (SNP) that was identified in exon 16. The variant allele was stable and was preferentially retained in RNA from the invasive ductal carcinoma of Case 72.

The difficulty in classifying the status of these large germline rearrangements highlights an area of increasing interest: noncoding variation in the human genome. The degree of variability within the *BRCA1* region has important implications for assessing the age of founder mutations. Additionally, an elevated mutation rate within the *BRCA1* region could implicate *BRCA1* as a target of genomic instability. It will be important to assess the degree and nature of noncoding variation within the *BRCA1* region in order to determine the effect of rare alleles such as the variant allele retained in the infiltrating ductal carcinoma of Case 72. The determination of the amount and the degree to which individual genomes vary will be critical to determining both the relevance to human disease for variant alleles such as the rare allele described from Case 72 and the use of genetic variation to infer the evolutionary history of human populations.

**Table 1 Restriction Fragment Length Polymorphisms (RFLP's) Identified in
BRCA1 from Breast Cancer Cases**

<u>Location</u>	<u>Primer Pair</u>	<u>Restriction Site</u>	<u>GenBank L78833^a</u>	<u>Frequency^b</u>
ivs 3	3F/4R	loss of Ava I	17,320	5/184 chromosomes
ivs 3	3F/4R	loss of Sty I	15,679	1/184 chromosomes
ivs 8	8F/11R	loss of Eco RI	30,718	7/184 chromosomes

- a. Location of the restriction cut site affected unless the specific nucleotide alteration was characterized.
b. The number of chromosomes in which the RFLP was identified.

**Table 2 Restriction Fragment Length Polymorphisms (RFLP's) Identified in
BRCA1 from Breast Cancer Cases**

<u>Location</u>	<u>Primer Pair</u>	<u>Restriction Site</u>	<u>GenBank L78833^a</u>	<u>Frequency^b</u>
ivs 3	3F/4R	loss of Ava I	17,320	5/184 chromosomes
ivs 3	3F/4R	loss of Sty I	15,679	1/184 chromosomes
ivs 8	8F/11R	loss of Eco RI	30,718	7/184 chromosomes

- a. Location of the restriction cut site affected unless the specific nucleotide alteration was characterized.
b. The number of chromosomes in which the RFLP was identified.

Figure 1

BRCA1 region of Chromosome 17q21

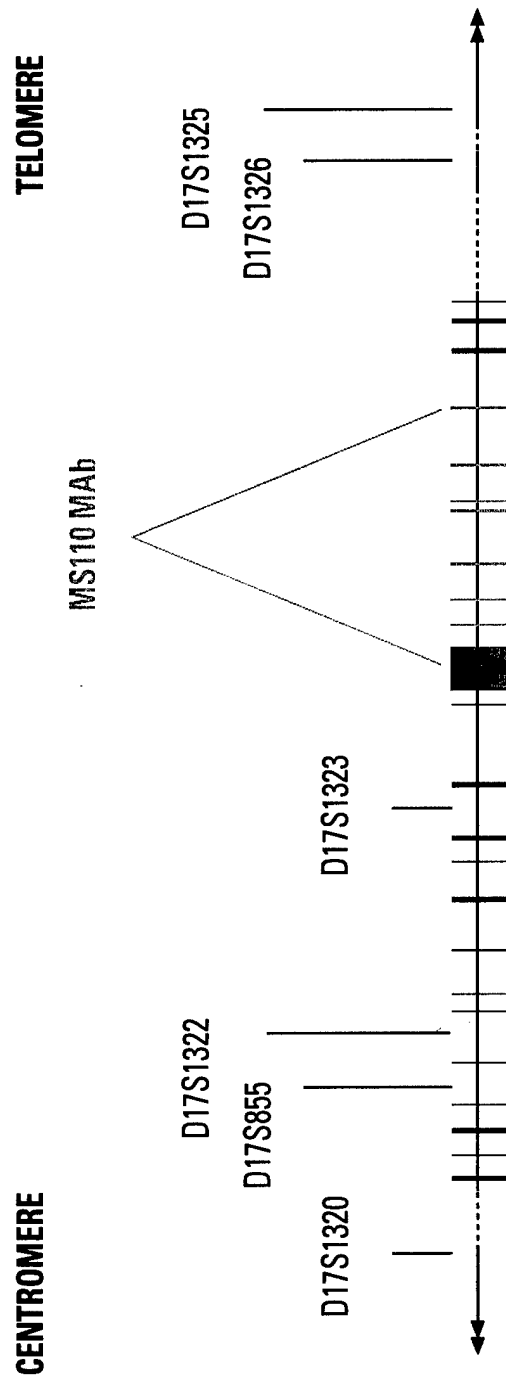
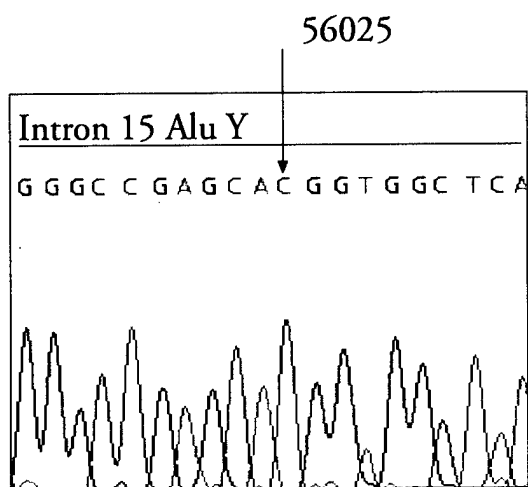


Figure 1 92 breast carcinomas were analyzed for LOH using the six microsatellite markers indicated in blue. 84 of 92 breast carcinomas were analyzed by Long PCR and 46 of 92 breast carcinomas were analyzed by Southern for large rearrangement of *BRCA1*. 68 of 92 breast carcinomas were analyzed by immunohistochemistry using the MS110 MAb. The *BRCA1* exons that encode the epitope recognized by MS110 are indicated in orange.

Figure 2

Wild-type BRCA1 Sequence:



977 bp Germline Deletion in Case 72:

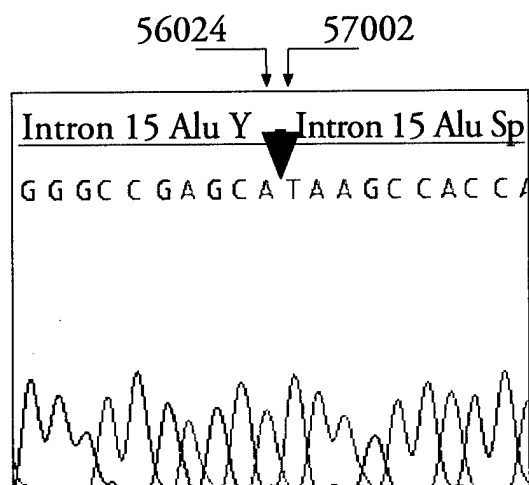


Figure 2 On the left is sequence from the unretained Case 72 *BRCA1* allele. The sequence from the unretained allele matches the GenBank L78833 sequence. On the right is sequence from the 977 bp deletion breakpoint junction of Case 72.

Figure 3

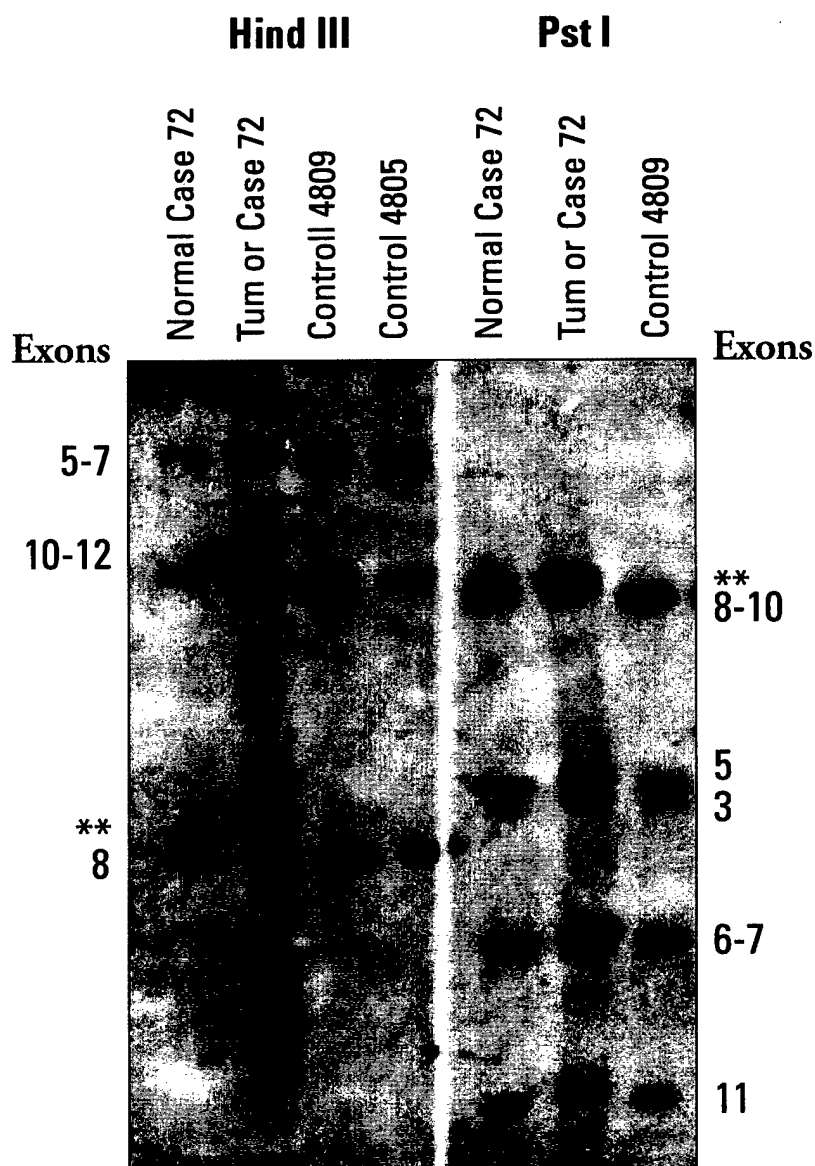
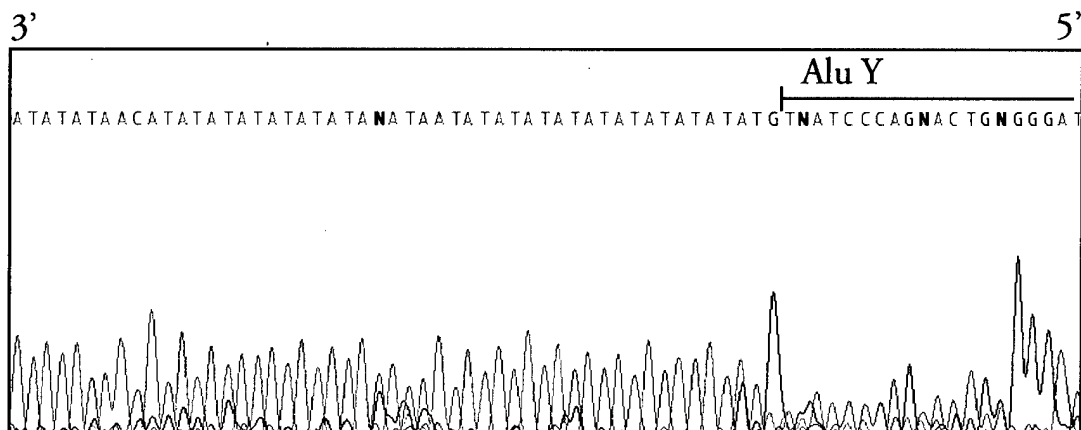


Figure 3 Southern analysis of a large germline rearrangement involving the *BRCA1* exon 8 genomic region. 10 μ g of restriction enzyme-digested genomic DNA was electrophoresed and Southern blotted as described in Materials and Methods. The blot was hybridized to a radioactively labelled probe made from exons 3 to 11 of the *BRCA1* cDNA. ** indicates the rearrangement junction fragment. Control individuals are from Family 48.

Figure 4

Wild-type BRCA1:



Case 72 Insertion:

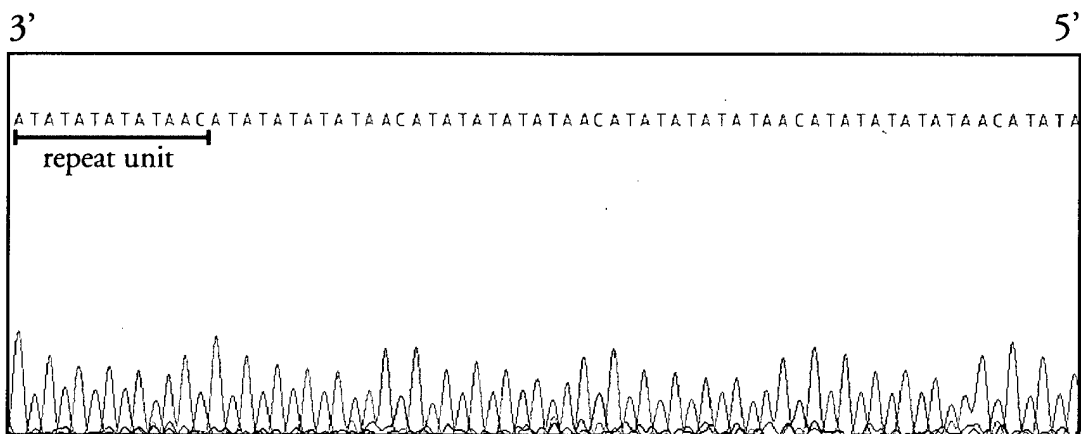


Figure 4 *BRCA1* intron 7 sequence from a control individual (4310) is shown in the top panel. The sequence from the control allele matches the GenBank L78833 sequence. Sequence from the *BRCA1* intron 7 insertion of ~250 bp identified in the retained allele from Case 72 is shown in the bottom panel. The insertion consists of an expansion of the compound microsatellite repeat unit (AT)_n AAC. Both sequences begin at approximately bp 28,084 of the GenBank L78833 sequence. The 3' and 5' symbols represent sequence orientation with respect to *BRCA1* transcription. *BRCA1* exon 8 is located at bp 28,853 to 28,957 of GenBank L78833. The compound microsatellite repeat unit is indicated by a red bar. The 3' tail of an Alu Y element is indicated in the control sequence. In the retained allele from Case 72, the Alu Y element lies ~250 bp downstream of the sequence shown.

Figure 5

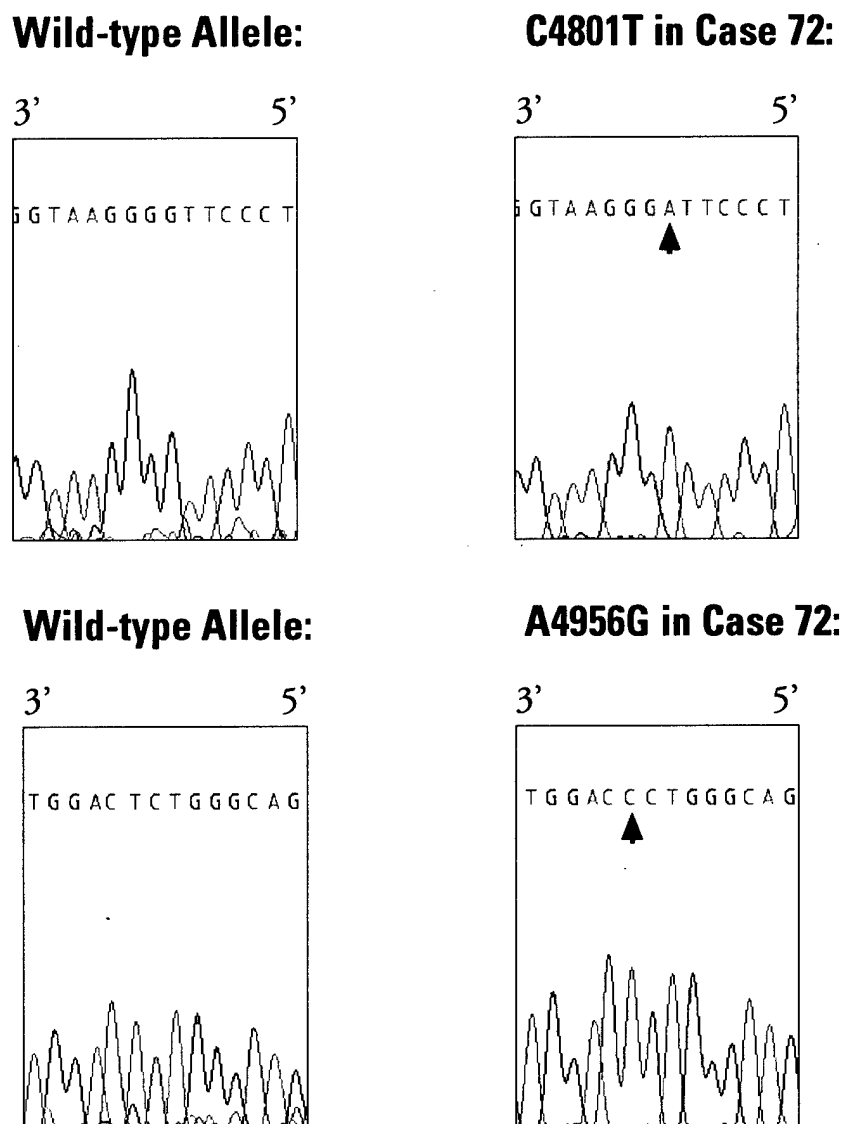
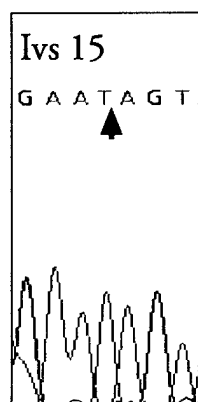


Figure 5 Single nucleotide polymorphisms identified in the variant allele from Case 72. The panels on the left represent wild-type sequence in the wild-type allele from Case 72. The panels on the right represent sequence of single nucleotide polymorphisms in *BRCA1* exon 16 of the variant allele from Case 72. In the top panels, sequence flanking a C to T transition at nt 4801 of HSU14680 is shown. In the bottom panels, sequence flanking a A to G transition at nt 4956 of HSU 14680 is shown. The red arrows indicate the location of the transition substitutions.

Figure 6

Wild-type:



24 bp Insertion in Case 72:

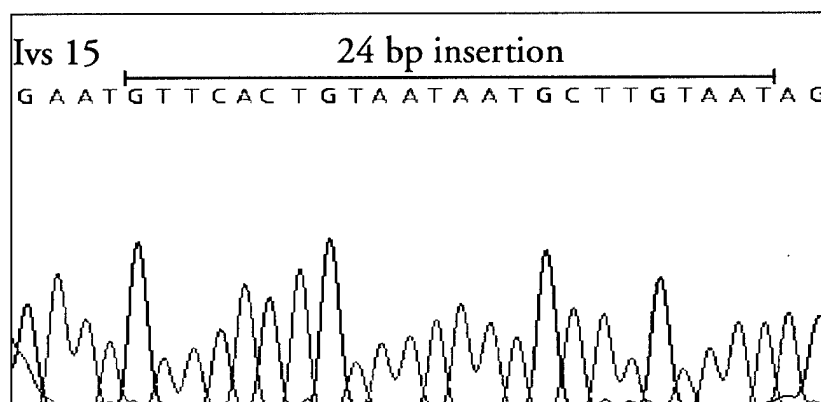


Figure 6 24 bp insertion in *BRCA1* intron 15. Sequence from the unretained Case 72 *BRCA1* allele is shown in the left panel. The sequence from the unretained allele matches the GenBank L78833 sequence. Sequence from the insertion junction fragment of the 24 bp identified in the variant allele from Case 72 is shown in the right panel. The red bar indicates the inserted nucleotides and the red arrow indicates the site of insertion.

References

- (1) Merajver SD, Pham TM, Caduff RF, Chen M, Poy EL, Cooney KA, Weber BL, Collins FS, Johnston C and Frank TS. 1995. Somatic mutations in the *BRCA1* gene in sporadic ovarian tumours. *Nature Genet.*, 9:439-443.
- (2) Breast Cancer Information Core. 1999.
http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/BIC.
- (3) Hosking L, Trowsdale J, Nicolai H, Solomon E, Foulkes W, Stamp G, Signer E and Jeffreys A. 1995. A somatic *BRCA1* mutation in an ovarian tumor. *Nature Genet.*, 9: 343-344.
- (4) Berchuck A, Heron K-A, Carney ME, Lancaster JM, Fraser EG, Vinson VL, Deffenbaugh AM, Miron A, Marks JR, Futreal PA and Frank TS. 1998. Frequency of germline and somatic *BRCA1* mutations in ovarian cancer. *Clin. Cancer Res*, 4: 2433-2437.
- (5) Khoo U-S, Ozcelik H, Cheung ANY, Chow LWC, Ngan HYS, Done SJ, Liang ACT, Chan VWY, Au GKH, Hg W-F, Poon CSP, Leung Y-F, Loong F, Ip P, Chan GSW, Andrulis IL, Lu J and Ho FCS. 1999. Somatic mutations in the *BRCA1* gene in Chinese sporadic breast and ovarian cancer. *Oncogene*, 18: 4643-4646.
- (6) Puget N, Torchard D, Serova-Sinilkova OM, Lynch HT, Feunteun J, Lenoir GM, Mazoyer, S. 1997. A 1-kb Alu-mediated germ-line deletion removing *BRCA1* exon 17. *Cancer Res*, 57: 828-831.
- (7) Swensen J, Hoffman M, Skolnick MH, Neuhausen SL. 1997. Identification of a 14 kb deletion involving the promoter region of *BRCA1* in a breast cancer family. *Hum Mol Genet*, 6: 1513-1515.
- (8) Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drüsedau M, Hogervorst FBL, Hageman S, Arts PJW, Ligtenberg MJL, Meijers-Heijboer H, Klijn, JGM, Vasen HFA, Cornelisse CJ, van't Veer LJ, Egbert-Bakker, van Ommen G-J.B, Devilee P. 1997. *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. *Nature Genet*, 17: 341-345.
- (9) Puget N, Stoppa-Lyonnet D, Sinilkova OM, Pages S, Lynch HT, Lenoir GM, Mazoyer S. 1999a. Screening for germ-line rearrangements and regulatory mutations in *BRCA1* led to the identification of four new deletions. *Cancer Res*, 59: 455-461.
- (10) Puget N, Sinilkova OM, Stoppa-Lyonnet D, Audoynaud C, Pages S, Lynch HT, Goldgar D, Lenoir GM, Mazoyer, S. 1999b. An alu-mediated 6-kb duplication in the *BRCA1* gene: a new founder mutation? *Am J Hum Genet*, 64: 300-302.
- (11) Montagna M, Santacatterina M, Torri A, Menin C, Zullato D, Chieco-Bianchi L and D'Andrea E. 1999. Identification of a 3 kb Alu-mediated *BRCA1* gene rearrangement in two breast/ovarian cancer families. *Oncogene*, 18: 4160-4165.
- (12) Smith SA, Easton DF, Evans DG and Ponder BA. 1992. Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nature Genet.*, 2: 128-131.
- (13) Purandare SM and Patel PI. 1997. Recombination hotspots and human disease. *Genome Res*. 7: 773-786.
- (14) Mazzarella R and Schlessinger D. 1998. Pathological consequences of sequence duplications in the human genome. *Genome Res*, 8: 1007-1021.

- (15) Brown MA, Xu CF, Nicolai H, Griffiths B, Chambers JA, Black D and Solomon E. 1996. The 5' end of the BRCA1 gene lies within a duplicated region of human chromosome 17q21. *Oncogene*, 12: 2507-2513.
- (16) Barker DF, Liu X and Almeida ER. 1996. The BRCA1 and 1A1.3B promoters are parallel elements of a genomic duplication at 17q21. *Genomics*, 38: 215-222.
- (17) Xu, C-F 1997. Complex regulation of the BRCA1 gene. *J Biol Chem*, 272: 20994-20997.
- (18) Newman B, Moorman PG, Millikan R, Qaqish BF, Geradts J, Aldrich TE, Liu ET. 1995. The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. *Breast Cancer Res Treat*, 35:51-60.
- (19) Newman B, Mu H, Butler LM, Millikan RC, Moorman PG and King M-C. 1998. Frequency of breast cancer attributable to BRCA1 in a population-based series of American women. *JAMA*, 279: 915-921.
- (20) Beckmann MW, Picard F, An HX, van Roeyen CR, Dominik SI, Mosny DS, Schnurch HG, Bender HG and Niederacher D. 1996. Clinical impact of detection of loss of heterozygosity of BRCA1 and BRCA2 markers in sporadic breast cancer. *Br J Cancer*, 73: 1220-6.
- (21) Niederacher D, Picard F, van Roeyen C, An HX, Bender HG and Beckmann MW. 1997. Patterns of allelic loss on chromosome 17 in sporadic breast carcinomas detected by fluorescent-labeled microsatellite analysis. *Genes Chromosom Cancer* 18:181-192.
- (22) Rio PG, Pernin D, Bay JO, Albuissou E, Kwiatkowski F, De Latour M, Bernard-Gallon DJ and Bignon YJ. 1998. Loss of heterozygosity of BRCA1, BRCA2 and ATM genes in sporadic invasive ductal breast carcinoma. *Int J Oncol*, 13: 849-853.
- (23) Silva JM, Gonzalez R, Provencio M, Dominguez G, Garcia JM, Gallego I, Palacios J, Espana P and Bonilla F. 1999. Loss of heterozygosity in BRCA1 and BRCA2 markers and high-grade malignancy in breast cancer. *Breast Cancer Res Treat*. 53: 9-17.
- (24) Dietmaier W, Wallinger S, Bocker T, Kullmann F, Fishel R, and Ruschoff J. 1997. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. *Cancer Res*, 57: 4749-56.
- (25) Wilson CA, Ramos L, Villasenor MR, Anders KH, Press MF, Clarke K, Karlan B, Chen JJ, Scully R, Livingston D, Zuch RH, Kanter MH, Cohen S, Calzone F.J and Slamon DJ. 1999. Localization of human BRCA1 and its loss in high-grade, non-inherited breast carcinomas. *Nature Genet.*, 21: 236-240.
- (26) Antonarakis SE and the Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. *Human Mutation*, 11:1-3.
- (27) Dunning AM, Chiano M, Smith NR, Dearden J, Gore M, Oakes S, Wilson C, Stratton M, Peto J, Easton D, Clayton D and Ponder BAJ. 1997. Common BRCA1 variants and susceptibility to breast and ovarian cancer in the general population. *Hum Mol Genet*, 6: 285-289.

Key Accomplishments

- 1) We designed a Long PCR screen for detecting large genomic rearrangement of *BRCA1*. The screen has been used successfully to identify two novel large germline rearrangements in our lab. A third large germline rearrangement of *BRCA1* was identified using Southern analysis.
- 2) We identified a complex large germline deletion within *BRCA1* that results in exon 3 being skipped in a family with inherited breast and ovarian cancer. The mutation differs from previous reports in that neither of the deletion breakpoints occurs in an Alu element and there is no more than 4 bp identity between sequence near the 5' breakpoint and sequence near the 3' breakpoint. Thus, this mutation is likely the result of a nonhomologous mechanism and is unique among large genomic rearrangements of *BRCA1*.
- 3) We analyzed 92 breast carcinomas unselected for family history using the Long PCR screen in combination with Southern analysis. Large somatic rearrangements of *BRCA1*, as with small mutations in *BRCA1*, are rare in sporadic breast cancer.
- 4) Reduced *BRCA1* protein expression was associated with genomic loss at D17S1323, the intragenic marker closest to the *BRCA1* transcription start site. The association of reduced *BRCA1* protein with *BRCA1* LOH in tumors is specific to LOH at marker D17S1323, the intragenic marker closest to the 5' end of *BRCA1*.
- 5) We characterized a rare variant allele of *BRCA1* in which two large germline rearrangements, as well as numerous small genomic alterations, reside in *cis*. Although the variant allele is retained in an invasive ductal carcinoma, no effect on the stability or splicing of *BRCA1* from the variant allele was observed. The difficulty in classifying the status of these large germline rearrangements highlights an area of increasing importance: noncoding variation in the human genome. The determination of the amount and the degree to which individual genomes vary will be critical to determining both the relevance to human disease for variant alleles such as the rare allele described from Case 72 and the use of genetic variation to infer the evolutionary history of human populations.

Key Accomplishments (cont.)

- 6) Many breast tumors in our series (55%, 12/22) had reduced BRCA1 protein staining despite no detectable genomic loss of the *BRCA1* transcription start as indicated by marker D17S1323. Reduced BRCA1 protein was observed in a high proportion (8/14, 57%) of the tumors that were informative at all three intragenic markers and fully retained both alleles of *BRCA1*. In this subset of tumors, other mechanisms are likely responsible for BRCA1 inactivation.

Reportable Outcomes

Abstracts:

Payne SR & King M-C. 1997. Complex BRCA1 germline deletion and rearrangement of 1039 bp leads to deletion of exon 3 and disease predisposition: Splicing by exon definition in BRCA1. *Am J Hum Genet*, 61 *Suppl*: A425.

Mu H, Newman B, Rousseau C, Payne S, Gordon K, Gold D, Tse C-K, Millikan RC & King M-C. 1999. Frequency of breast cancer attributable to BRCA1 and BRCA2 in a population-based series of caucasian and African-American women. *Am J Hum Genet*, 65 *Suppl*: A106.

Manuscripts:

Payne SR, Newman B & King M-C. Complex germline rearrangements of BRCA1 associated with breast and ovarian cancer. *Genes, Chromosomes & Cancer*, accepted.

Presentations:

Payne SR, Gonzalez-Hernandez & King M-C. 1999. BRCA1 in sporadic breast cancer: Genomic loss and protein expression in primary tumors. *Am J Hum Genet*, 65 *Suppl*: A109.

Appendix One: Published Abstracts

Payne SR & King M-C. 1997. Complex BRCA1 germline deletion and rearrangement of 1039 bp leads to deletion of exon 3 and disease predisposition: Splicing by exon definition in BRCA1. *Am J Hum Genet*, 61 Suppl: A425.

BRCA1 mutations detectable by current PCR-based techniques are single bp substitutions or small insertions/deletions which comprise most germline mutations in high-risk families, but almost never occur as somatic mutations. Conversely, Mbp somatic deletions including BRCA1, as reflected by LOH, occur frequently in both inherited and sporadic breast and ovarian cancer. We are characterizing an intermediate class of mutations, involving deletions or rearrangements of 100's to 1000's of bp. One such germline mutation has been reported in BRCA1 (Feunteun et al., 1997. *Cancer Res*, 57:828).

In Family 5, previous analysis of cDNA revealed germline deletion of exon 3 (codon 27 stop), cosegregating with 10 cases of breast and ovarian cancer (Friedman et al., 1994. *Nat Genet.*, 8:399); the genomic basis of the mRNA deletion remained unknown. Long-range PCR, using an exon 3 forward/"exon 4" reverse primer pair designed from BRCA1 genomic sequence (gb L78833), reveals a 1039 bp genomic deletion involving part of exon 3 and part of intron 3. The mutant genomic sequence is: wt bp1-22 of exon 3; insert inverted bp 22-13 of exon 3; deletion of 1039 bp; wt bp 14015 ff intron 3. Despite wt 5' and 3' splice sites for intron 2, exon 3 is completely skipped. This mutation indicates that at least in some regions of BRCA1, splicing follows the exon definition model. Next, we will determine whether intermediate length deletions occur as somatic mutations in sporadic breast and ovarian tumors.

Appendix One: Published Abstracts (cont.)

Mu H, Newman B, Rousseau C, Payne S, Gordon K, Gold D, Tse C-K, Millikan RC & King M-C. 1999. Frequency of breast cancer attributable to BRCA1 and BRCA2 in a population-based series of caucasian and African-American women. *Am J Hum Genet*, 65 Suppl: A106.

To determine the public health impact of inherited mutation in BRCA1 or BRCA2 on breast cancer, both genes were fully screened in a population-based series of sequentially diagnosed Caucasian and African-American breast cancer patients from North Carolina. New results for BRCA2 and recently discovered large genomic alterations complete our previously published results on BRCA1 in the same cohort (Newman et al., *JAMA* 279: 915-921, 1998). BRCA2 was analyzed for germline variants in the coding sequence, splicing junctions and neighboring intronic regions using multiplex single-strand conformation analysis, heteroduplex analysis, and DNA sequencing of 203 cases. In addition, 373 cases were screened for the large genomic mutations. After adjustment for sampling probabilities, the weighted prevalences (%) of breast cancer attributable to BRCA1 or BRCA2 were 7.4 (95% CI, 1.3-12.1) for Caucasian women, 1.0 (95% CI, 0-3.0) for African-American women, and 6.1 (95% CI, 1.2-9.8) overall. Among probands with at least three affected relatives, 13% carried a BRCA1 or BRCA2 mutation. Among breast cancer probands with any relative with ovarian cancer, 22% carried a mutation. Among American breast cancer patients generally, those from these high risk families can most benefit from full genotyping of BRCA1 and BRCA2.

Appendix One: Published Abstracts (cont.)

Payne SR, Gonzalez-Hernandez & King M-C. 1999. BRCA1 in sporadic breast cancer: Genomic loss and protein expression in primary tumors. *Am J Hum Genet*, 65 Suppl: A109.

In inherited breast cancer, BRCA1 acts as a classic tumor suppressor in that both normal copies of BRCA1 are lost, one in the germline and one somatically. In sporadic tumors, large somatic deletions including BRCA1, as reflected by loss of heterozygosity (LOH), occur frequently, while somatic point mutations are rare. One argument for the role of BRCA1 in sporadic breast cancer has been the high rate of allelic loss observed for the BRCA1 locus. Sporadic invasive breast tumors were evaluated for protein expression and genomic loss at BRCA1. Protein expression was evaluated by immunohistochemistry using the N-terminal monoclonal antibody MS110. Genomic loss was evaluated by LOH using multiple markers within and flanking BRCA1, by Long PCR (7-12kb fragments) of the entire BRCA1 genomic region and by Southern analysis. BRCA1 protein expression was significantly decreased in 72% (49/68) of invasive ductal carcinomas and correlated significantly with tumor grade, as previously observed (Wilson et al, 1999 *Nat Genet* 21:236, Lee et al, 1999 *Histopath* 34:106). No large genomic rearrangements of somatic origin were detected, although one germline rearrangement was detected by Southern analysis. Although allelic loss at all intragenic markers is not associated with BRCA1 protein reduction, allelic loss at the marker nearest the transcription start site (D17S1323) is associated with reduction of BRCA1 protein. This observation might reflect specific targeting of either the BRCA1 promoter or BRCA1 exon 11 for loss. The function of BRCA1 that is involved in maintenance of genetic stability through the G2-M checkpoint has been shown to require exon 11 (Xu et al, 1999 *Mol Cell* 3:389). A high proportion (57%, 8/14) of the tumors that retained both BRCA1 alleles showed loss of the BRCA1 protein. In this subset of tumors other mechanisms may be responsible for BRCA1 inactivation.

Appendix Two: Manuscripts

**Complex germline rearrangement of BRCA1 associated with breast and
ovarian cancer.**

Shannon R. Payne¹, Beth Newman^{2,3}, and Mary-Claire King⁴

¹ Department of Genetics, University of Washington, Seattle, WA, USA 98195

² Lineberger Cancer Center, University of North Carolina, Chapel Hill, North Carolina
27599

³ Current Address: School of Public Health, Queensland University of Technology,
Brisbane, Queensland, 4059, Australia

⁴ Departments of Genetics & Medicine, Division of Medical Genetics, University of
Washington, Seattle, WA, USA 98195

Corresponding Author: Mary-Claire King

Telephone: 206-616-4294

Fax: 206-616-4295

e-mail: mcking@u.washington.edu

Running Title: Complex rearrangement of BRCA1

ABSTRACT

Germline mutations in *BRCA1* predispose to breast and ovarian cancer. Most germline *BRCA1* mutations are small insertions, deletions or single base pair (bp) substitutions. These mutation classes are rarely found as somatic mutations in *BRCA1*. Conversely, somatic deletions of multiple megabase pairs (Mb) including *BRCA1*, as reflected by loss of heterozygosity, occur frequently in both inherited and sporadic breast and ovarian cancer. In order to determine whether deletions or rearrangements of hundreds to thousands of bp might contribute to inherited mutation in *BRCA1*, we developed a Long PCR strategy for screening the entire genomic *BRCA1* locus in high-risk families. We evaluated genomic DNA from one high-risk family of Western European ancestry with *BRCA1*-linked cancer in which no genomic mutations had been detected using conventional methods. Long PCR revealed a complex mutation, g.12977 ins10 del 1039 (based on GenBank L78833) comprising an inverted duplication and deletion in *BRCA1* that removes portions of exon 3 and intron 3, including the 5' splice site for intron 3. As a result of the deletion exon 3 is skipped, leading to a truncated protein and disease predisposition. Unlike previously reported large germline deletions in *BRCA1*, neither breakpoint resides within an Alu element. The g.12977 ins10 del1039 mutation was not detected among eleven other breast cancer families, nor among 406 breast cancer patients unselected for family history.

Keywords: *BRCA1*, large deletion, RNA splicing, exon skipping

Germline protein-truncating mutations in *BRCA1* cause hereditary predisposition to breast and ovarian cancer (Hall et al., 1990; Ford et al., 1994). *BRCA1* mutations detectable by commonly employed PCR-based techniques are small insertions, deletions or single base pair (bp) substitutions which comprise most germline mutations in high-risk families (Breast Cancer Information Core. 1999), but rarely occur as somatic mutations (Futreal et al., 1994; Hosking et al., 1995; Matsushima et al., 1995; Merajver et al., 1995; Takahashi et al., 1995; Berchuck et al., 1998). Conversely, megabasepair (Mb) somatic deletions including *BRCA1*, as reflected by loss of heterozygosity (LOH), occur frequently in both inherited and sporadic breast and ovarian cancer (Devilee et al., 1994; Bieche et al., 1995). We have developed a protocol to screen for an intermediate class of mutations, involving deletions or rearrangements of hundreds to thousands of bp. Several large deletions and one large duplication involving Alu repetitive elements have been reported previously as germline mutations in *BRCA1* (Puget et al., 1997; Swensen et al., 1997; Petrij-Bosch et al., 1997; Puget et al., 1999a; Puget et al., 1999b).

We report here the identification and characterization of a large genomic deletion in a family with inherited breast and ovarian cancer. The mutation results in skipping of exon 3 in the mature *BRCA1* message. Although the genomic mutation was not detectable using conventional PCR-based *BRCA1* screening procedures, it was identified using a Long PCR strategy designed to detect intermediate length genomic deletions and rearrangements.

Previous analysis of lymphoblast cDNA from Family 5 revealed a transcript lacking exon 3 in several *BRCA1*-linked family members (Friedman et al., 1994; Fig.1). Deletion of exon 3 maintains the reading frame, but creates a stop at codon 27. Splice junctions of exons 2, 3, and 5 (the exon joined to the 3' splice site of exon 3) were sequenced from PCR products using primers immediately flanking the exons, yielding only wild-type sequence (Friedman et al., 1994). The genomic basis of the variant remained unknown.

For the analysis described here, gDNA from family members was amplified using the fifteen PCR primer pairs indicated in Table 1. Primers were designed from *BRCA1* genomic sequence (GenBank L78833). Primer pairs were designed to amplify 3 to 11 kb genomic fragments with at least 1 kb of overlap. PCR products were evaluated for differences from predicted size before and after digestion with appropriate restriction enzymes.

Using primer pair 3 designed to amplify a 5531 bp product encompassing exon 3, a variant Long PCR product co-segregated with the *BRCA1*-linked haplotype of Family 5 (Fig. 1, 2). All family members heterozygous for transcripts lacking exon 3 were also heterozygous for the variant PCR product.

Wild-type and variant Long PCR products were cloned and sequenced from multiple family members. Sequence of all variant clones revealed the complex mutation shown in Figure 3. Based on *BRCA1* genomic sequence L78833, the mutation is a 10bp inverted duplication of 12965-12974 from *BRCA1* exon 3 (bp 210-219 of *BRCA1* cDNA, HSU14680) and deletion of 1039bp. The notation for the mutation (Antonarakis et al., 1998) is g.12977 ins10 del1039 based on GenBank L78833. The net deletion is 1029 bp, yielding an amplified product of 4502 bp which is consistent with the electrophoretic mobility of the variant PCR product. The 1039 bp genomic deletion in Family 5 results in skipping of *BRCA1* exon 3 in the mRNA and premature protein truncation at codon 27.

Three other families with breast cancer linked to *BRCA1* were tested for the g.12977 ins10 del1039 mutation. None of the families contained the g.12977 ins10 del1039 mutation. Families with multiple cases of breast and or ovarian cancer and with negative LOD scores for linkage to *BRCA1* might nonetheless harbor mutations in one of these genes if some cases are sporadic (ie. phenocopies). Seven such families were screened for g.12977 ins10 del1039. All were wild-type at this site. Finally, we investigated the contribution of the g.12977 ins10 del1039 mutation to breast cancer in a population-based

series of 242 white and 164 African-American breast cancer patients unselected for family history. Patients were ascertained previously as part of the Carolina Breast Cancer Study (Newman et al., 1995; Newman et al., 1998). A PCR primer pair was designed to detect a 514 bp breakpoint junction fragment for the g.12977 ins10 del1039 mutation. Primers used to amplify the breakpoint junction fragment were: (forward) 5'-TTT-TTC-TCC-CCC-CCT-ACC-CTG-3'; (reverse) 5'-GCT-CAG-CAT-TTG-TTA-CTC-AAG-CTG-3'. No mutation carriers were detected in 406 cases.

The mutation in Family 5 differs from previously reported large germline deletions in *BRCA1*. Many large *BRCA1* mutations involve Alu sequences (Puget et al., 1997; Swensen et al., 1997; Petrij-Bosch et al., 1997; Puget et al., 1999a; Puget et al., 1999b). For two of these, both the 5' and 3' breakpoints reside in Alu sequence. In one of these families, a 1 kb deletion of exon 17 is mediated by oppositely oriented Alu elements, removing most parts of both Alu elements from the mutant allele (Puget et al., 1997). In another family, a 14 kb deletion involving the *BRCA1* promoter region creates a new Alu element (Swensen et al., 1997). In contrast, the 5' breakpoint for the Family 5 mutation occurs within *BRCA1* exon 3 and the 3' breakpoint occurs in intron 3 at position +1008. Neither of these breakpoints occurs in an Alu element. Further, there is no more than 4 bp identity between sequence near the 5' breakpoint and sequence near the 3' breakpoint, suggesting that this mutation is the result of a nonhomologous event.

Combinations of inversions and deletions have been seen in both somatic and germline mutations of several genes including *TP53* (Greenblatt et al., 1996) and *Factor IX* (Ketterling et al., 1994; Sommer, 1995). Studies of somatic mutations in *HPRT* reveal alterations with interesting similarity to the naturally occurring *BRCA1* germline mutation of Family 5. Somatic deletions in *HPRT* often include inserted bases at the breakpoint junctions that are inverted complements of sequences found at the breakpoints (Rainville et al., 1995). In *HPRT*, topoisomerase I and topoisomerase II have been implicated in

generating the free DNA ends that may result in these nonhomologous recombination events (Rainville et al., 1995; Monnat et al., 1992). Topoisomerase II sites also occur near rearrangement breakpoints in the *Dystrophin* gene (Hu et al., 1991) and ring chromosome 21 (Wong et al., 1989). *In vitro* and *in vivo* evidence for the role of topoisomerase II in nonhomologous recombination is well established in prokaryotes (O'Connor et al., 1985; Ikeda, 1986) and a role for vertebrate topoisomerase II in nonhomologous recombination has been demonstrated *in vitro* (Bae et al., 1988). Topoisomerase II boxes are in close proximity to the 3' deletion breakpoint of the complex mutation in Family 5. Sequences matching 13 of 15 bp of the topoisomerase II d box, GTN(A/T)A(T/C)ATTNATNN(A/G), occur in *BRCA1* intron 3 (Sander and Hsieh, 1985). Either the sequence tTAAGTATTTATGTA at position +948 or the sequence aTGTACATTTtTCTG at +1225 might serve as a recognition site for cleaving the double stranded DNA at +1008 of *BRCA1* intron 3.

Mutations in *BRCA1* and *BRCA2* account for at least 80% of hereditary breast and ovarian cancer (Narod et al., 1995; Rebbeck et al., 1997). Complex *BRCA1* and *BRCA2* mutations such as that in Family 5 account for a still unknown fraction of unexplained families with inherited breast and ovarian cancer (Narod et al., 1995; Rebbeck et al., 1997; Schubert et al., 1997). Additionally, there may be other as yet uncharacterized genes that predispose to hereditary breast cancer (Rebbeck et al., 1997; Schubert et al., 1997).

The mutation in Family 5 adds to a growing class of mutations not detectable by conventional screening methods. Amplification followed by genomic SSCP analysis would invariably amplify the wild-type allele in Family 5. Direct sequencing of splice junctions would not identify the genomic mutation because the primer routinely used to analyze the 5' splice site for intron 3 was deleted in the variant allele. Reports of genomic deletions of *BRCA1* that were undetected by conventional methods further illustrate the limitations in standard *BRCA1* and *BRCA2* screening procedures.

As more large genomic deletions and rearrangements are characterized in the *BRCA1* region, several questions remain to be addressed. Are Alu-Alu recombination events prominent in the *BRCA1* region, as seen for deletions in the β -globin gene cluster (Henthorn et al., 1990), or no more likely to be involved than random sequences in the region? Likewise, which is more representative of large deletions and rearrangements in the *BRCA1* region: the complex inverted duplication with deletion observed in Family 5 in this study or the simple deletions reported previously? Of primary importance is whether such mutations occur somatically as well as in the germline. If such mutations do exist somatically, it will be interesting to determine whether they are similar in Alu-dependence and complexity to those observed in the germline.

ACKNOWLEDGEMENTS

We wish to thank Jan Morrow for technical assistance, Piri Welch, Elizabeth Schubert, and Raymond Monnat for critical comments and discussion, and Roddy Grant for assistance with figures. This work was supported in part by NIH grant RO1-CA27632. Shannon Payne is supported in part by NIH training grant T32-GMO7735.

REFERENCES

- Antonarakis SE and the Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. *Human Mutation* 11:1-3.
- Bae YS, Kawasaki I, Ikeda H, Liu LF. 1988. Illegitimate recombination mediated by calf thymus DNA topoisomerase II *in vitro*. *Proc Natl Acad Sci USA* 85: 2076-2080.
- Berchuck A, Heron K-A, Carney ME, Lancaster JM, Fraser EG, Vinson VL, Deffenbaugh AM, Miron A, Marks JR, Futreal PA, Frank TS. 1998. Frequency of germline and somatic *BRCA1* mutations in ovarian cancer. *Clin Cancer Res* 4:2433-2437.
- Bieche I, Lidereau R. 1995. Genetic alterations in breast cancer. *Genes Chromosomes Cancer* 14: 227-251.
- Breast Cancer Information Core. 1999.
http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/BIC
- Devilee P, Cornelisse CJ. 1994. Somatic genetic changes in human breast cancer. *Biochem Biophys Acta* 1198: 113-130.
- Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE and the Breast Cancer Linkage Consortium. 1994. Risks of cancer in *BRCA1*-mutation carriers. *Lancet* 343:692-695.
- Friedman LS, Ostermeyer EA, Szabo CI, Dowd P, Lynch E, Rowell SE, King, M-C. 1994. Confirmation of *BRCA1* by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nature Genet* 8:399-404.
- Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y, Eddington K, McClure M, Frye C, Weaver-Feldhaus J, Ding W, Gholami Z, Söderkvist P, Terry L, Skolnick MH, Kamb A, Wiseman R. 1994. *BRCA1* mutations in primary breast and ovarian carcinomas. *Science* 266:120-122.
- Greenblatt MS, Grollman AP, Harris CC. 1996. Deletions and insertions in the p53 tumor

suppressor gene in human cancers: Confirmation of the DNA polymerase slippage/misalignment model. *Cancer Res* 56: 2130-2136.

Hall JM, Lee MK, Newman B, Morrow J, Anderson LA, Huey B, King, M-C. 1990.

Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250:1684-1689.

Henthorn PS, Smithies O, Mager DL. 1990. Molecular analysis of deletions in the human β -globin gene cluster: Deletion junctions and locations of breakpoints. *Genomics* 6: 226-237.

Hosking L, Trowsdale J, Nicolai H, Solomon E, Foulkes W, Stamp G, Signer E, Jeffreys

A. 1995. A somatic *BRCA1* mutation in an ovarian tumor. *Nature Genet* 9: 343-344.

Hu X, Ray PN, Worton RG. 1991. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J* 10:2471-2477.

Ikeda H. 1986. Bacteriophage T4 DNA topoisomerase mediates illegitimate recombination *in vitro*. *Proc Natl Acad Sci USA* 83: 922-926.

Ketterling RP, Vielhaber EL, Lind TJ, Thorland EC, Sommer SS. 1994. The rates and patterns of deletions in the human Factor IX gene. *Am J Hum Genet* 54: 201-213.

Matsushima M, Kobayashi K, Emi M, Saito H, Saito J, Suimori K, Nakamura Y. 1995. Mutation analysis of the *BRCA1* gene in 76 Japanese ovarian cancer patients: four germline mutations, but no evidence of somatic mutation. *Hum Mol Genet* 4:1953-1956.

Merajver SD, Pham TM, Caduff RF, Chen M, Poy EL, Cooney KA, Weber BL, Collins FS, Johnston C, Frank, T.S. 1995. Somatic mutations in the *BRCA1* gene in sporadic ovarian tumours. *Nature Genet* 9: 439-443.

Monnat RJ, Hackmann AFM, Chiaverotti TA. 1992. Nucleotide sequence analysis of human hypoxanthine phosphoribosyltransferase (HPRT) gene deletions. *Genomics* 13: 777-787.

- Narod SA, Ford D, Devilee P, Barkardottir RB, Lynch HT, Smith SA, Ponder BJ, Weber BL, Garber JE, Birch JM, Cornelis RS, Kelsell DP, Spurr NK, Smyth E, Haites N, Sobol H, Bignon Y-J, Chang-Claude J, Hamann U, Lindblom A, Borg A, Piver S, Gallion HH, Struwing JP, Whittemore A, Tonin P, Goldgar DE, Easton DF, the Breast Cancer Linkage Consortium. 1995. An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. *Am J Hum Genet* 56: 254-264.
- Newman B, Moorman PG, Millikan R, Qaqish BF, Geradts J, Aldrich TE, Liu ET. 1995. The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. *Breast Cancer Res Treat* 35:51-60.
- Newman B, Mu H, Butler L, Millikan RC, Moorman PG, King M-C. 1998. Frequency breast cancer attributable to *BRCA1* in a population-based series of American women. *JAMA*. 279:915-921.
- O'Connor MB, Malamy MH. 1985. Mapping of DNA gyrase cleavage sites *in vivo*: oxolinic acid induced cleavages in plasmid pBR322. *J Mol Biol* 181: 545-550.
- Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drüsedau M, Hogervorst FBL, Hageman S, Arts PJW, Ligtenberg MJL, Meijers-Heijboer H, Klijn, JGM, Vasen HFA, Cornelisse CJ, van't Veer LJ, Egbert-Bakker, van Ommen G-J.B, Devilee P. 1997. *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. *Nature Genet* 17: 341-345.
- Puget N, Torchard D, Serova-Sinilkova OM, Lynch HT, Feunteun J, Lenoir GM, Mazoyer, S. 1997. A 1-kb Alu-mediated germ-line deletion removing *BRCA1* exon 17. *Cancer Res* 57: 828-831.
- Puget N, Stoppa-Lyonnet D, Sinilkova OM, Pages S, Lynch HT, Lenoir GM, Mazoyer S. 1999a. Screening for germ-line rearrangements and regulatory mutations in *BRCA1* led to the identification of four new deletions. *Cancer Res* 59: 455-461.
- Puget N, Sinilkova OM, Stoppa-Lyonnet D, Audouyoud C, Pages S, Lynch HT, Goldgar D, Lenoir GM, Mazoyer, S. 1999b. An alu-mediated 6-kb duplication in the *BRCA1*

- gene: a new founder mutation? *Am J Hum Genet* 64: 300-302.
- Rainville IR, Albertini RJ, Nicklas JA. 1995. Breakpoints and junctional regions of intragenic deletions in the HPRT gene in human T-cells. *Som Cell Mol Genet* 21: 309-326.
- Rebbeck TR, Couch FJ, Kant J, DeShano M, Peng Y, Chen K, Garber JE, Weber BL. 1997. Genetic heterogeneity in hereditary breast cancer: Role of BRCA1 and BRCA2. *Am J Hum Genet* 60:547-553.
- Sander M, Hsieh T-S. 1985. Drosophila topoisomerase II double-strand DNA cleavage: analysis of DNA sequence homology at the cleavage site. *Nuc Acids Res* 13: 1057-1072.
- Schubert EL, Lee MK, Mefford HC, Argonza RH, Morrow JE, Hull J, Dann JL, King, M-C. 1997. BRCA2 in American families with four or more cases of breast or ovarian cancer: Recurrent and novel mutations, variable expression, penetrance, and the possibility of families whose cancer is not attributable to BRCA1 or BRCA2. *Am J Hum Genet* 60: 1031-1040.
- Sommer SS. 1995. Recent human germline mutation: Inferences from patients with hemophilia B. *Trends Genet* 11: 141-147.
- Swensen J, Hoffman M, Skolnick MH, Neuhausen SL. 1997. Identification of a 14 kb deletion involving the promoter region of BRCA1 in a breast cancer family. *Hum Mol Genet* 6: 1513-1515.
- Takahashi H, Behbakht K, McGovern PE, Chiu H-C, Couch FJ, Weber B, Friedman LS, King M-C, Furusato M, LiVlosi VA, Menzin AW, Liu PC, Benjamin I, Morgan MA, King SA, Rebane BA, Cardonick A, Mikuta JJ, Rubin SC, Boyd J. 1995. Mutation analysis of the *BRCA1* gene in ovarian cancers. *Cancer Res* 55: 2998-3002.
- Wong C, Kazazian HH, Stetten G, Earnshaw WC, Van Keuren ML, Antonarakis SE. 1989. Molecular mechanism in the formation of a human ring chromosome 21. *Proc Natl Acad Sci USA* 86: 1914-1918.

Legends to Figures

Figure 1. Pedigree for Family 5. Three generation pedigree for Family 5 indicating markers used to establish linkage. Below the linkage markers are displayed the results of both SSCP analysis of cDNA and Long PCR analysis of genomic DNA using primer pair no. 3 designed to amplify a genomic region including *BRCA1* exon 3. Some unaffected family members are not included.

Figure 2. Variant allele reflecting *BRCA1* g.12977 ins10 del1039 in Family 5 detected by Long PCR of genomic DNA. Numbers above lanes are identification numbers from the pedigree. The wild-type genomic product extends from *BRCA1* intron 2 to intron 3 and is 5531 bp. Amplification from the mutant allele yields a 4502 bp product. The variant band is present in all individuals with the *BRCA1*-linked haplotype and who are heterozygous for deletion of exon 3 in *BRCA1* mRNA.

Figure 3. Exon skipping resulting in a truncated BRCA1 protein in Family 5.

(A) Wild-type BRCA1 sequence and splicing of exon 3. nt 210-240 of HSU14680 shown.

(B) Breakpoint sequence of the 1039 bp deletion and effects on splicing in Family 5.

nt 210-221 of HSU14680 and 10 bp inverted duplication, followed by intron 3 sequence.

Table 1: Long PCR primers used for analysis of BRCA1 gene.

Pair No. ^a	Primer Name	Temp. ^b	Size (bp)	GenBank L78833 ^c	Sequence
1	n2.1.1bR BA2	65	2364	2309-4672	5'-GCA ATG CAA AGA CCG TCC GCT G-3' 5'-GTA CTT CTT CAA CGC GAA GAG CAG ATA AAT C-3'
2	1aF 3R	65	9936	3028-12963	5'-TTT GGA CAA TAG GTA GCG ATT CTG ACC TTC-3' 5'-AAC TCC AGA CTA GCA GGG TAG GGG GGG-3'
3	3F 4R	61	5531	12792-18322	5'-TCC TGA CAC AGC AGA CAT TTA-3' 5'-CCC GTC TCT ACA GAA AAC AC-3'
4	ivs3F 7R	65	10857	13715-24571	5'-TGT GAA GAC AGG AAA GGA CCT GAT ACC AGT TTC-3' 5'-CAC GGT TTC TGT AGC CCA TAC TTT GGA TGA TAG-3'
5	6F-2 ivs7R	65	3828	23839-27666	5'-GCT TTT CAG CTT GAC ACA GGT TTG G-3' 5'-CCC CAG CAC TCC TAA GAA CAT TTA GTA TAG G-3'
6	8F 11R	65	8361	28853-37213	5'-CAG GAA ACC AGT CTC AGT GTC CAA CTC TCT AAC CTT G-3' 5'-TGT CAC TCA GAC CAA CTC CCT GGC TTT CAG AC-3'
7	11F 13R	65	9518	36783-46300	5'-CCA TAC ACA TTT GGC TCA GGG TTA CCG AAG AGG G-3' 5'-TTC GCA GGT CCT CAA GGG CAG AAG AGT CAC-3'
8	12F ivs13R	65	8927	37715-46641	5'-GGT GTG AGA GTG AAA CAA GCG TCT CTG AAG ACT GC-3' 5'-GCC TGT CAC CAA TTT CTC CCA TTC CAC TTA GCT TC-3'
9	13F 15R	65	8071	46235-54275	5'-GGA GCC AGC CTT CTA ACA GCT ACC CTT CCA TC-3' 5'-GAC TCC CAG AGC AAC TGT GCA TGT ACC ACC TAT C-3'
10	15F-2 18R-2	65	10613	54241-64852	5'-TGA TAG GTG GTA CAT GCA CAG TTG CTC TGG-3' 5'-GCT AAC TAC CCA TTT TCC TCC CGC AAT TCC-3'
11	17F 19R	61	4581	60922-65502	5'-GTG TAG AAC GTG CAG GAT TG-3' 5'-CAT TGT TAA GGA AAG TGG TGC-3'
12	18F 20R	65	6897	64777-71673	5'-TGC AGA TGC TGA GTT TGT GTG TGA ACG GAC-3' 5'-CCT GGG ATT CTC TTG CTC GCT TTG GAC C-3'
13	ivs19F 21R	65	7495	70154-77648	5'-TCC CAG TGA GGT GAA AAG CCG ATT GTT AAG TTC-3' 5'-CCC ATA GCA ACA GAT TTC TAG CCC CCT GAG G-3'
14	20F 23R	61	9662	71518-81179	5'-ATA TGA CGT GTC TGC TCC AC-3' 5'-ACT GTG CTA CTC AAG CAC CA-3'
15	22F ivs24R	63	8536	79560-88095	5'-AGC TGT GTG GTG CTT CTG TGG TGA AGG-3' 5'-AGA GCC AGC AAG ATC AGA TGG TCT ACA GGA C-3'

a. Long PCR was carried out using the Boehringer Mannheim Expand Long Template PCR System

b. Temp. = Annealing Temperature

c. Region of GenBankL78833 amplified by primer pair

Figure 1

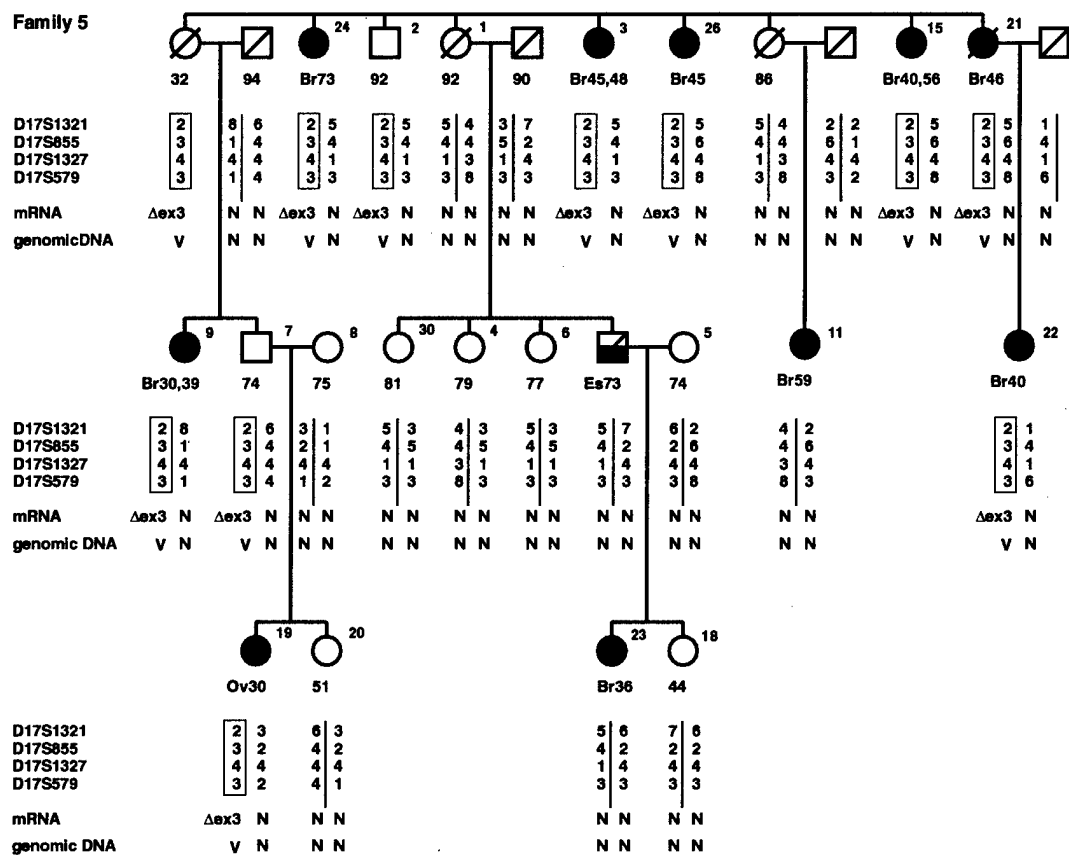


Figure 2

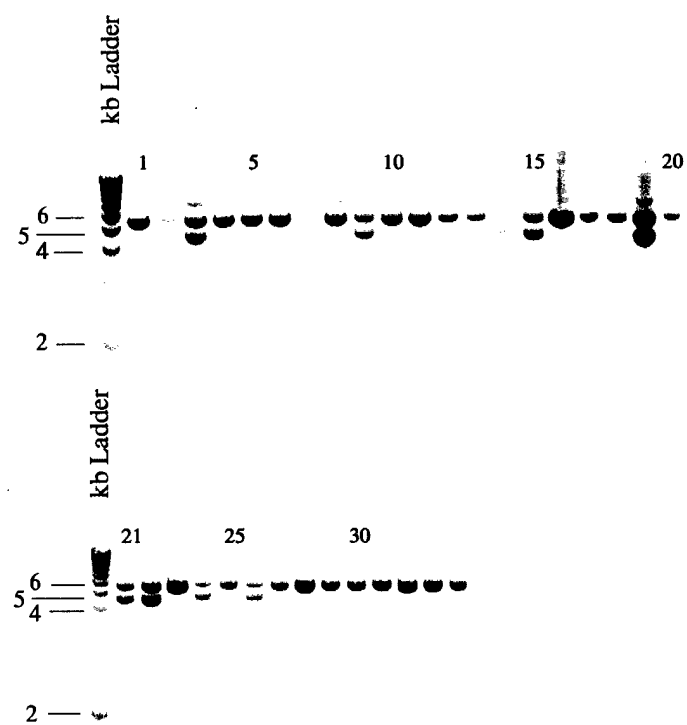
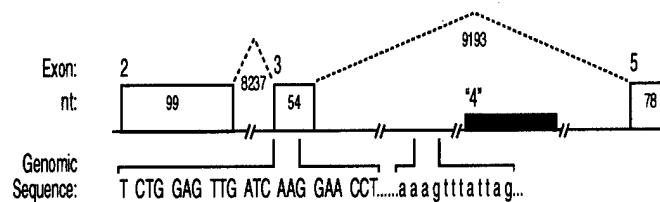
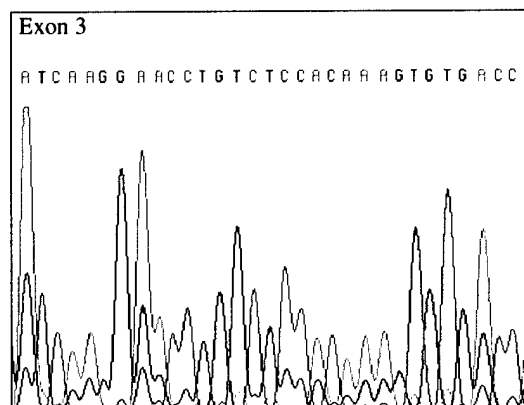


Figure 3

A. Wild-type BRCA1 Sequence:



B. 1039 bp Deletion in Family 5:

